DOCUMENT RESUME

ED 174 681 TM 009 653

AUTHOR Estes, Gary D.
TITLE Estimating Gains and Establishing Program Objectives Using Data from Nationally Normed Tests.
PUB DATE Apr 79
NOTE 18p.; Paper presented at the Annual Meeting of the American Educational Research Association (63rd, San Francisco, California, April 8-12, 1979)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Achievement Gains; Achievement Tests; Compensatory Education Programs; *Educational Objectives; Elementary Secondary Education; Evaluation Methods; Models; *National Norms; *Norm Referenced Tests; Predictive Measurement; *Program Evaluation; Scores
IDENTIFIERS California Achievement Tests; Comprehensive Tests of Basic Skills; *Elementary Secondary Education Act Title I; Normal Curve Equivalent Scores; SRA Assessment Survey

ABSTRACT

Normative data reported by publishers of the California Achievement Test, the Comprehensive Tests of Basic Skills, and the SRA Assessment Survey were used to illustrate a procedure for specifying goals for Title I programs in terms of normal curve equivalent (NCE) gains. The procedure was based on the amount of growth reflected in student reading and mathematics norms from one grade level to another, grades two to eleven inclusive. Specifically, the spring-to-spring standardized gains used as the no-treatment expectation were calculated, by subtracting the pretest means at the lower grade level from the posttest means at that grade, and dividing by the standard deviation on the pretest. Administrators can use the procedure, presented as tables of data, to subjectively judge Title I impact as a function of the percentage of increase which would be expected without the program. The procedure should be used with caution, however, because NCE gains across levels were not comparable, and because gains at higher grade levels were increasingly difficult to observe. (CP)

ESTIMATING GAINS AND ESTABLISHING PROGRAM OBJECTIVES USING DATA FROM NATIONALLY NORMED TESTS

Gary D. Estes

Northwest Regional Educational Laboratory

Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 1979.

ESTIMATING GAINS AND ESTABLISHING PROGRAM OBJECTIVES USING DATA FROM NATIONALLY NORMED TESTS

Introduction

A frequent question which arises in relation to the proposed Title I evaluation models and reporting system is, "How many NCEs can we expect or should we specify for our project objective(s)?" One response is that it is too early in the use of the models and NCE scores to estimate the size of NCE gains which might be expected. Another response offered is that "all NCE gains greater than zero are good!" (Tallmadge, 1976). Neither of these responses is likely to give project personnel the information they desire in order to assess the relative success achieved with Title I-funded programs.

Data reported by many test publishers might be useful to a decision maker or project manager in determining the size of NCE gain to specify as an objective. Specifically, the data for nationally normed tests provide evidence of the year-to-year gains for students in the norm sample. Data from three commonly used nationally normed tests were used in outlining a procedure to specify outcome objectives for Title I programs in terms of NCE gains.

The procedure which follows allows the project director or others to specify the gain for Title I students as a function of a percentage increase in the achievement which normally would be expected, i.e., without the Title I program. The gains can be specified by using data from one or more tests and within or across grade levels. It is not

necessary to use a nationally normed test for evaluating the program in order to specify the desired NCE gain using the proposed procedure.

The procedure uses as a basis the amount of growth reflected in student norms from one grade level to another. For example, if the mean of the third grade norms is one-half a standard deviation above the mean of the second grade norms, it would be reasonable to conclude that a gain of .5 standard deviations can be expected without any special assistance. If some special assistance were provided which was expected to increase student achievement by 50 percent, then a gain of .75 standard deviations would be expected.

The procedure uses student gains calculated using expanded standard scores as the basis for calculating NCE gains. The gain in expanded standard score units from a "pretest" to "posttest" for the normed group is divided by the standard deviation on the pretest to estimate the expected gain or "no-treatment expectation" in standard deviation units. Tallmadge and Fagan (1977) used a similar procedure but used the standard deviation at grade 3.7 to rescale data across all grades. Project objectives can be expressed as a percent of the no-treatment expectations and converted to the metric of Normal Curve Equivalents (NCEs) by multiplying the resultant product by 21.06. These calculations can be expressed as follows:

No-Treatment Expectation = (Posttest Mean - Pretest Mean) - Pretest Standard Deviation;

where:

means and standard deviations are in expanded standard score metrics.

NCE Treatment Effect = Percent x No-Treatment Expectation x 21.06;

where:

percent is the extra growth expected as a result of the special program expressed as a proportion of the normal growth.

The procedure outlined in this paper assumes that within the normed group there is a common gain across the different levels, i.e., students +1, 0 or -1 standard deviations from the norm will remain in the same relative position; that is, equipercentile growth is assumed. To the extent that lower level students fall farther behind in standard deviation units, the less useful these data will be for special populations, e.g., Title I programs. If the "falling off" is observed, the procedure will need to be applied separately for students within various ranges in the norm samples. Data provided by Tallmadge and Fagan (1977) and Stenner, Hunter, Bland and Cooper (1978) are evidence that standardized growth rates for different percentile levels are different across grade levels.

Procedure and Results

Data from the California Achievement Tests (CAT), 1970, Comprehensive Tests of Basic Skills (CTBS), 1974, and SRA Assessment Survey (SRA), 1974, were used to illustrate the procedure for estimating NCE gains. The expanded standard score means and standard deviations for grades 2 through 11 on the CAT Reading and Math Computation subtests (CAT, p. 14), the CTBS Total Reading and Math Computation subtests (CTBS, 1974, pp. 34-37), and the SRA Reading and Math tests (SRA, 1974, p. 23) are reported in Tables 1-3. The spring-to-spring standardized gains used as the "no-treatment" expectation or normal growth expectation were calculated for each grade by subtracting the "pretest" means at the lower grade level from the "posttest" means at that grade and dividing by the

standard deviation on the pretest. The sample objectives were 50 percent of the no-treatment standardized gains. The NCE gains used as sample objectives were then obtained by multiplying the objectives in standard deviation units by 21.06.

In Table 1, the .85 for third grade implies that from the spring of the second grade to the spring of the third grade the normed group gained .85 second grade standard deviations and that from the spring of the third grade to the spring of the fourth grade the norm group moved .49 third grade standard deviations.

The data in Tables 1-3 demonstrate that not only do the variances in student achievement increase across grades but that, in general, the amounts of student growth expressed in standard deviation units decreases rapidly until the fifth or sixth grade and that some leveling occurs in the higher grades. These data are consistent with the findings of Stenner et al. (1978). Standard deviation gains for the reading trends in Figure 1 and the math trends in Figure 2 illustrate this relation.

These data are evidence that different gains should be expected across grade levels. Stenner et al. (1978) and Tallmadge and Fagan (1977) observed similar results. The implications are that gains will likely not be comparable across grade levels independent of the "quality" of programs and that it will be increasingly difficult to observe gains at higher grade levels due to the increase in variances and decrease in gains at the higher grades.

How might these data be useful in establishing the objectives? It is suggested that one could establish a project objective by examining the data for a test and calculating the "NCE" gain similar to the examples in

Tables 1-3. More than one test might be helpful if the intent were to balance out the effects of the relative gains among different test series. If data for more than one test were used, the average gain across tests could be used at each grade level, as was done by Tallmadge and Fagan (1977). However, based on the data for the three tests reported here, it appears that the gains from different test series might be fairly consistent.

The procedure illustrated above for establishing objectives involves making a subjective judgment about the amount of impact a Title I program would have above and beyond that which would be anticipated with the regular school program. For example, assume that a project was using the '70 CAT to evaluate a third grade reading and math program. According to the data in Table 1, gains of 9.1 and 17.7 NCEs would be expected for reading and math respectively, if it was felt that the Title I program could increase normal student achievement by 50 percent. However, if the decision maker felt that the Title I program should accelerate student achievement by 25 percent, then program objectives of 4.5, i.e., 25 x .85 x 21.06 = 4.5, and 8.8, i.e., .25 x 1.67 x 21.06 = 8.8, NCE gains would be appropriate for reading and math, respectively. If the decision maker felt that student achievement should be doubled, then program objectives of 17.9 and 35.1 NCE gains would be appropriate for reading and math, respectively. Given the ninth grade data in Table 1, it is apparent that an NCE gain of 3.2 would be an appropriate objective if it were felt that the Title I reading program should increase by 50 percent the amount which students normally would achieve without any Title I assistance. An

NCE gain of 1.5 would be an appropriate objective if it were felt that the ninth grade Title I program should boost student achievement by 25 percent.

The above example can be characterized as a subjective approach to establishing objectives since the objectives were determined on the basis of someone's value judgment about the amount of impact which would be expected. A more objective approach might be to use the proportion of supplementary funds expended on Title I students. That is, if 25 percent more funds were spent on Title I students than would have been spent without the Title I program, 25 percent might be used as the basis for establishing the amount of increase in learning expected as a result of the Title I program. This, of course, assumes a direct relationship between level of funding and program effectiveness which is likely to be very tenuous.

Also, it might be possible to use the amount of supplementary assistance Title I students receive. If a Title I program increased the amount of instruction Title I students received in reading by 25 percent, then it might be reasonable to use 25 percent as the estimated increase in reading attributable to Title I.

Conclusion

While the procedure outlined is best described as a rough guideline, some conclusions seem warranted. First, evidence was presented which implied that NCE gains across grade levels will probably not be comparable. This does not imply that one could not, or should not, aggregate data across grade levels. The analogy of measuring pounds of fruit cocktail by adding apples and bananas to describe the amount of fruit one has might apply to aggregating data across grade levels. What

will be questionable is to judge a 6 NCE gain in a third grade reading program and a 4 NCE gain in a ninth grade reading program as evidence that the third grade reading program was more successful.

The procedure outlined might provide an interim basis for stating program objectives until sufficient data are available for observing the size and range of NCE gains associated with programs in different subject areas and grade levels. It is likely that the evaluations which will be reported in the near future will provide empirical data as to the size of NCE gains which can be expected.

Table 1

CAT 1970

Expanded Standard Score Means and Standard Deviations
and Spring-to-Spring Gains, by Year,
for Reading and Math Computation

	READING					MATH				
Grade	Mean	SD	No-Treatment Expectation	Sample Objectives: 50% Increase[1] SDs	NCEs	Mean	SD	No-Treatment Expectation	Sample Objectives: 50% Increase[1] SDs	NCEs
2	327	53	--	--		282	27	--	--	
3	372	61	.85	.43	9.1	327	38	1.67	.84	17.7
4	402	65	.49	.25	5.3	363	44	.95	.48	10.1
5	440	72	.59	.30	6.3	414	60	1.16	1.08	22.7
6	465	78	.35	.18	3.8	458	74	.73	.37	7.8
7	500	86	.45	.23	4.8	497	91	.53	.27	5.7
8	531	92	.36	.18	3.8	535	101	.42	.21	4.4
9	558	95	.29	.15	3.2	565	102	.30	.15	3.2
10	600	100	.44	.22	4.6	600	100	.34	.17	3.6
11	630	103	.30	.15	3.2	621	108	.21	.11	2.3

[1]50% increase objective implies that Title I program would increase by one-half the amount of achievement expected without the program.

Table 2

CTBS 1974

Expanded Standard Score Means and Standard Deviations
and Spring-to-Spring Gains, by Year,
for Reading and Math Computation

	READING					MATH				
Grade	Mean	SD	No-Treatment Expectation	Sample Objectives: 50% Increase[1] SDs	NCEs	Mean	SD	No-Treatment Expectation	Sample Objectives: 50% Increase[1] SDs	NCEs
2	328	62.2	--	--		305	37.7	--	--	
3	380	70.8	.84	.42	8.8	361	47.5	1.48	.74	15.6
4	422	78.6	.59	.30	6.2	390	49.0	.61	.31	6.4
5	459	84.9	.47	.24	4.9	436	64.7	.94	.47	9.9
6	485	93.2	.31	.16	3.3	462	74.1	.40	.20	4.2
7	514	100.0	.31	.16	3.3	487	82.9	.34	.17	3.6
8	545	105.1	.31	.16	3.3	513	90.7	.31	.16	3.4
9	571	110.1	.25	.12	2.7	546	108.0	.36	.18	3.8
10	612	111.6	.37	.19	3.9	571	105.3	.23	.12	2.5
11	645	112.3	.30	.15	3.2	588	105.7	.16	.08	1.7

[1]50% increase objective implies that Title I program would increase by one-half the amount of achievement expected without the program.

Table 3

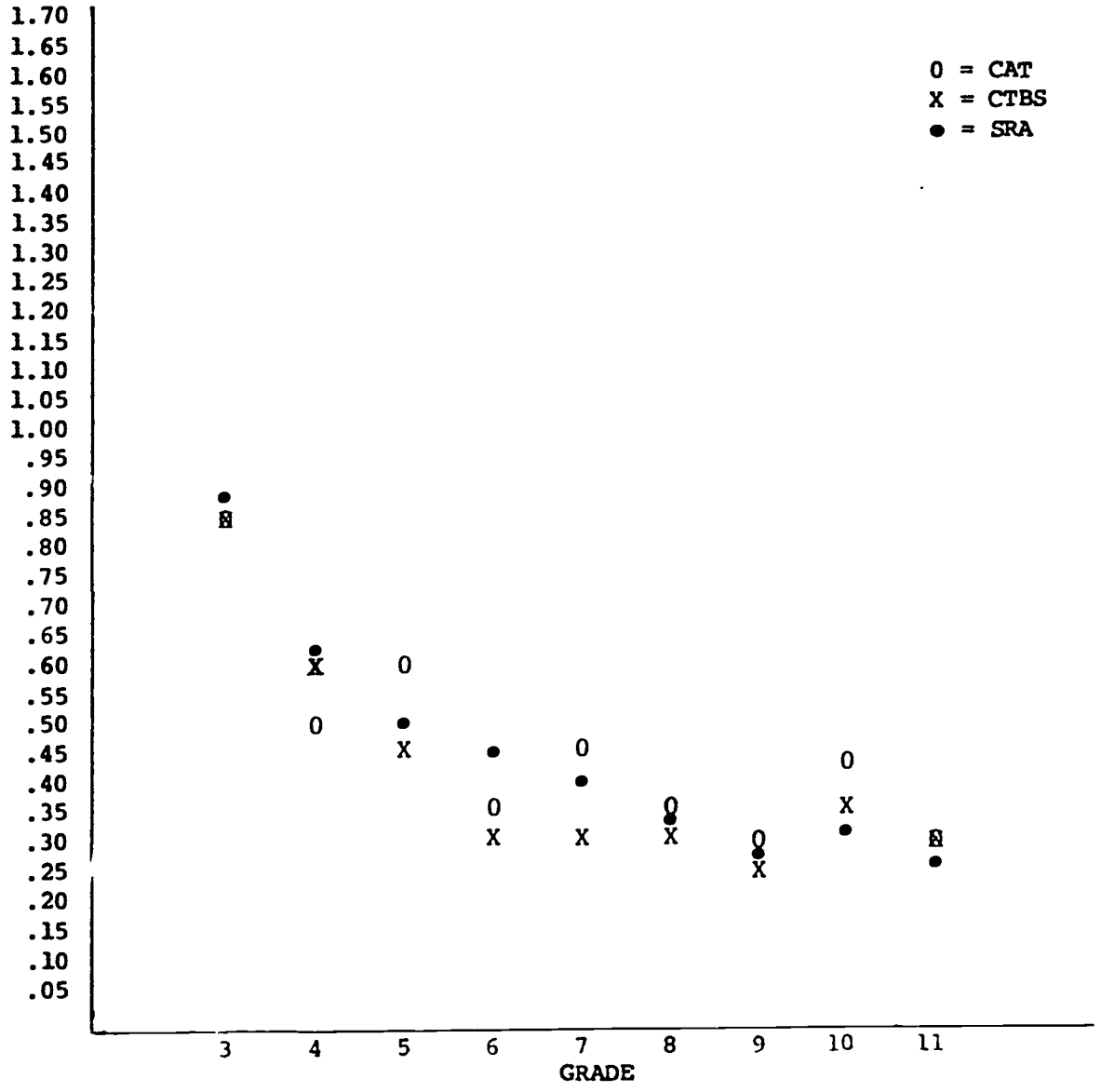
SRA 1974

Expanded Standard Score Means and Standard Deviations
and Spring-to-Spring Gains, by Year,
for Reading and Math Computation

	READING					MATH				
Grade	Mean	SD	No-Treatment Expectation	Sample Objectives: 50% Increase[1] SDs	NCEs	Mean	SD	No-Treatment Expectation	Sample Objectives: 50% Increase[1] SDs	NCEs
2	190	54	--	--		180	30	--	--	
3	236	58	.85	.43	9.1	232	41	1.73	--	
4	273	62	.64	.32	6.7	274	50	1.02	.51	10.7
5	304	62	.50	.25	5.3	307	56	.66	.33	6.9
6	332	62	.45	.23	4.8	346	67	.70	.35	7.4
7	356	62	.39	.20	4.2	376	74	.45	.23	4.8
8	337	62	.34	.17	3.6	402	80	.35	.18	3.8
9	394	62	.27	.14	2.9	420	82	.23	.12	2.5
10	414	63	.33	.17	3.6	443	87	.28	.14	2.9
11	433	66	.30	.15	3.2	465	95	.25	.13	2.7

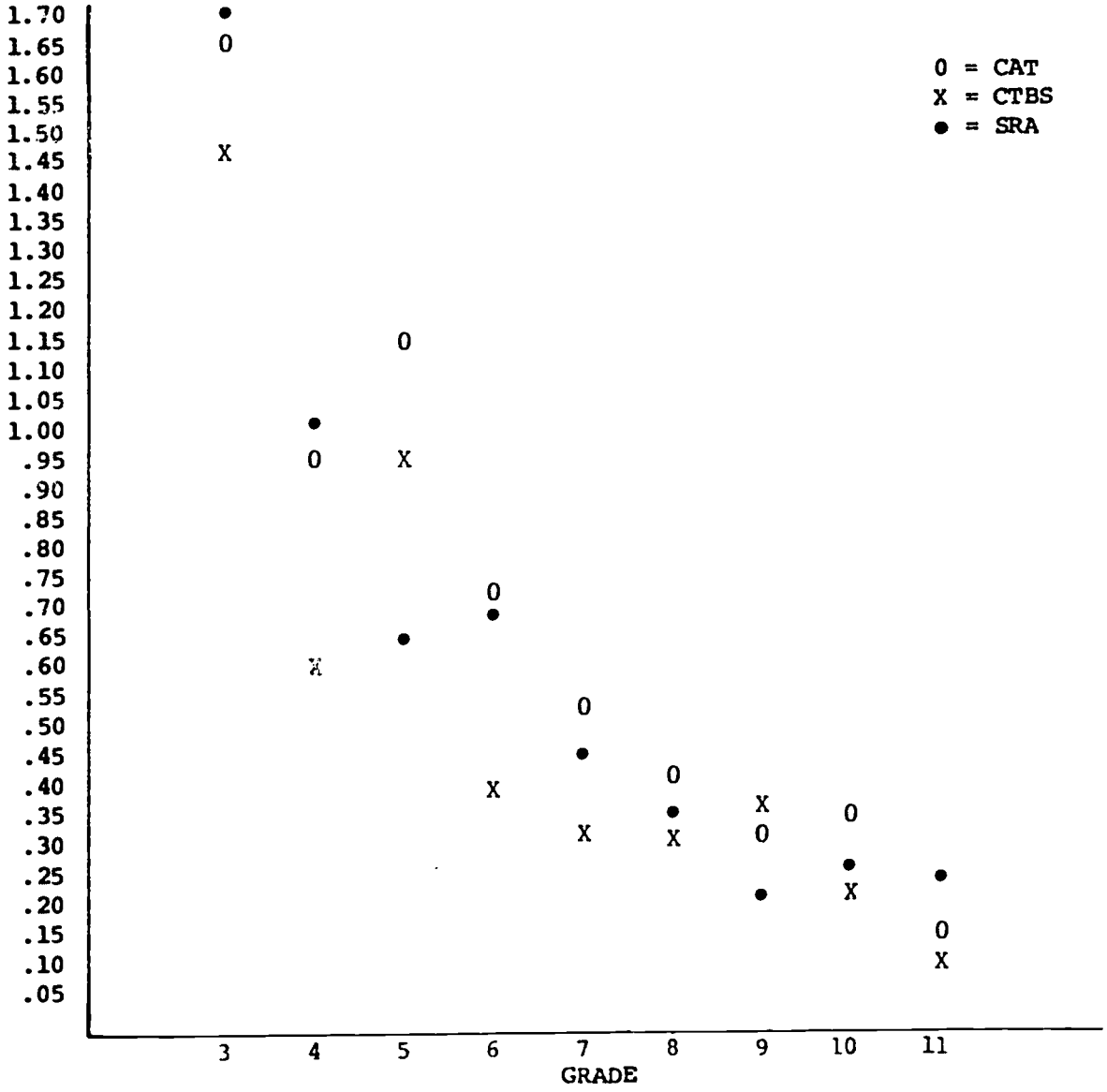
[1]50% increase objective implies that Title I program would increase by one-half the amount of achievement expected without the program.

FIGURE 1



Gains, Expressed in Standard Deviations
for CAT, CTBS and SRA Reading Subtests, by Grade

FIGURE 2



Gains, Expressed in Standard Deviations
for CAT, CTBS and SRA Reading Subtests, by Grade

REFERENCES

California Achievement Tests: Bulletin of Technical Data Number 1. Monterey, Ca.: CTB/McGraw Hill, 1971.

Comprehensive Tests of Basic Skills Form S: Technical Bulletin Number 1. Monterey, Ca.: CTB/McGraw Hill, 1974.

SRA Assessment Survey: Technical Report. Chicago: Science Research Associates, 1974.

Stenner, A. J., Hunter, E. L., Bland, J. D., & Cooper, M. L. The standardized growth expectation: Implications for educational evaluation. Paper presented at the meeting of the American Educational Research Association, Toronto, April 1978.

Tallmadge, G. K., & Fagan, B. M. Cognitive growth and growth expectations in reading and mathematics. Mountain View, Ca.: RMC Research Corporation, November 1977.